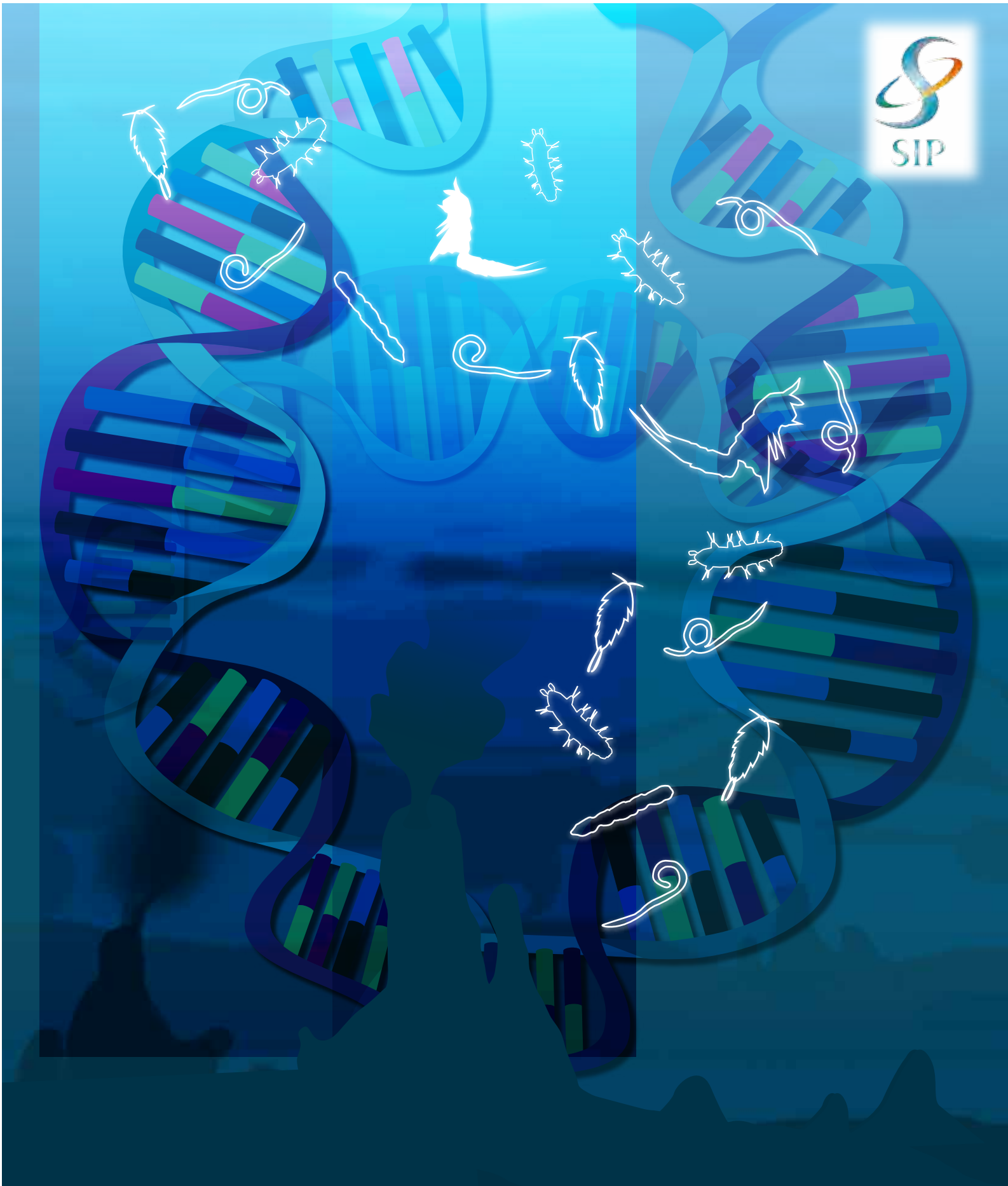
SIP Protocol Series No.1

Application of environmental metagenomic analyses for environmental impact assessments



Cross-ministerial Strategic Innovation Promotion Program

The Strategic Innovation Promotion Program (SIP) was launched by the Council for Science, Technology, and Innovation (CSTI), which oversees projects that target scientific and technological innovation in line with Japanese government directions as stated in the Comprehensive Strategy on Science Technology and Innovation and the Japan Revitalization Strategy. This interdisciplinary program among government agencies, academic institutes and private sectors addresses eleven issues. One of these issues is Next-Generation Technology for Ocean Resources Exploration.

Zipangu in the Ocean Program and Protocols for Environmental Survey Technologies

Zipangu in the Ocean Program is a technical study of the development of submarine mineral deposits that takes into consideration the wise use of these resources.

One research area is the ecological survey of organisms and their long-term monitoring. However, an ecosystem consists of various interrelated factors; thus, in addition to a comprehensive understanding of the system, observation and analysis of each component to its most elemental level are unavoidable. Recently, increased environmental awareness and the necessity of forming a consensus have become key issues in conducting development activities. Growing concern for the environment by the public and the diversification of the use of maritime areas have complicated the interests of stakeholders. To facilitate the formation of a consensus under these conditions, it is important for standardized methods to be implemented. This will ensure that research processes are transparent and that the collection of survey data is objective.

This protocol series aims to introduce more accurate, user-friendly, objective and effective underlying technologies required to understand the environmental impact of submarine mineral resource development. We believe that creating such a technology tool-kit will allow us to develop these resources in a sustainable manner.

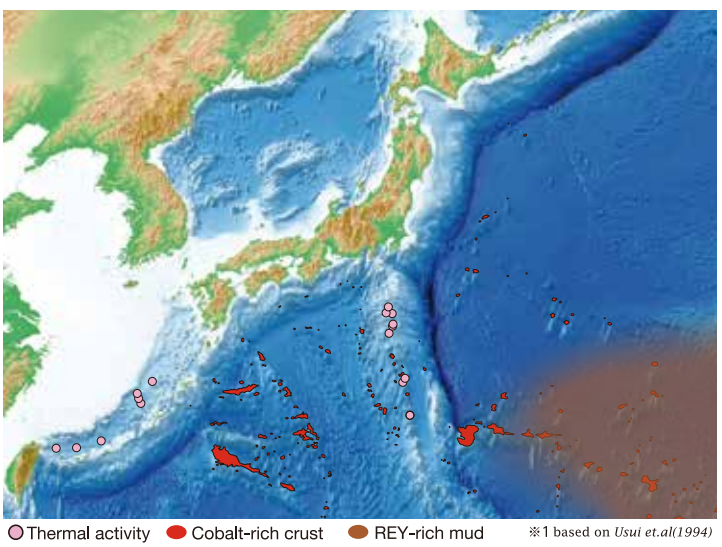


Table of Contents

chapter 1

Introduction

"In order to know the biological community in a particular environment, for example if the target is a plant (flora) it is necessary to collect plants, prepare specimens and identify / classify them. Moreover, from the number of individuals, it is also possible to grasp to what extent a certain organism lives in the area (environment) and the proportion of the species of the area.

However, when the target is microorganisms[1], microscopic observation and culturing of microbes are necessary. Therefore, the procedures are much more troublesome. Moreover, there are many microbes that cannot be cultured; therefore, it has not been possible to comprehensively study those organisms until recently. In addition, specific knowledge and skills may be required for identifying collected organisms in some cases, and the process is not simple for everybody.

In the recent rise in development of seabed mineral resources, there is a strict requirement for environmental impact assessment by the development. For example, identification of organisms at the species level may be required. Moreover, with this movement, it is required to investigate the biodiversity of the development area and to monitor changes in biodiversity (ISA, ISBA/19/LTC/8). However, regions are being developed where humans have little experience, such as deep seabed; thus, processing ecological information and environmental impact evaluation is unknown territory.

On the other hand, development in recent molecular biological methods has been remarkable; therefore, it became possible to study the biota of environmental samples such as water and soil by sequencing diverse DNAs extracted from the environmental samples, without purifying a specific organism from the environmental samples. The results will be affected by efficiency rates of each of the processes, such as the DNA extraction[2] and PCR amplification[3]; however, the number of sequences obtained are thought to be almost proportional to the number of organism species in the environment. Therefore, the quantitative levels are used to compare differences in the biological community in the samples.

Until recently, biological community analysis using diverse DNA has been performed by methods such as amplifying a target gene by PCR[4], cloning[5] of the amplified gene, and then the DGGE method (Denaturing Gradient Gel Electrophoresis)[6] where DNA derived from each organism is extracted and purified before sequencing analysis. Currently, gene sequences derived from organisms in diverse DNA can be directly determined by next generation sequencer and additionally it has become possible to analyze a large amount of nucleotide sequences by one analysis.

One of the processes in biological community analysis using the next generation sequencer (NGS) is called amplicon sequence analysis. In this process, first determine a target gene, the gene sequence was amplified by PCR, analyzing the sequences amplified by NGS (amplicons). Furthermore, shotgun sequencing where all of the gene sequences obtained in the diverse DNA are sequenced without amplifying a specific gene using PCR, is distinguished from amplicon sequence analysis. Nucleotide sequences derived from various genes can be obtained by shotgun sequencing; however, it is difficult to identify biological species where the nucleotide sequence is derived from. Amplicon sequence analysis and shotgun sequencing are both called metagenomics analysis; however, it should be noted that the concepts are different from one another. Furthermore, analysis targeting functional genes using RNA is called transcriptomic analysis. An analytical scheme for the biota in environmental samples using next generation sequencer described above is shown in Fig. 1

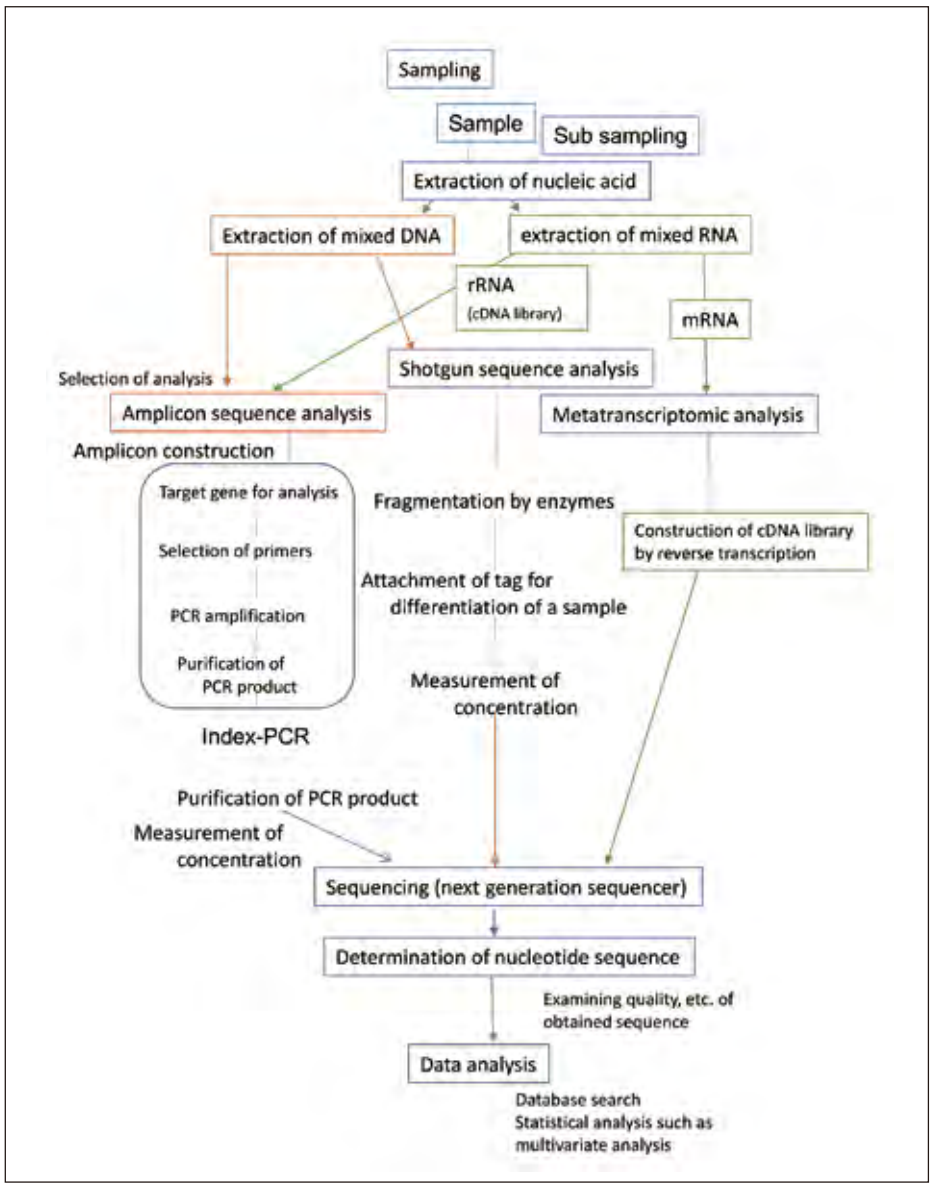


Fig. 1 Analytical scheme for the biota in environmental samples using next generation sequencer

chapter 2

Analysis of biocenosis by amplicon sequence analysis

Amplicon sequence analysis, one of analytical methods for the structure of biological community using next generation sequencer is described below. Main processes in amplicon sequence analysis are sampling, DNA extraction, amplicon construction by PCR, sequencing, and data analysis.

2-1. Sampling

Sampling can be said to be the most important factor determining the result of subsequent steps.

Sampling can be performed using various methods depending on purpose; however, it is important not to contaminate samples with exogenous organisms. In addition, it is necessary to determine the preservation method during transporting samples from a collection site to a laboratory to maintain samples at the best condition. Sediment should be collected by an adequate core sampler and a core (a cylinder shape sample, Fig. 2, left) is sliced from the top to create various layers for sampling (Fig. 2, right). Pictures are taken and morphology of each core (color, texture (condition of particles, etc.), other characters, etc.) is described as well as a collection site and method, a core number, position of the layer, analytical use, preservation method should be recorded.

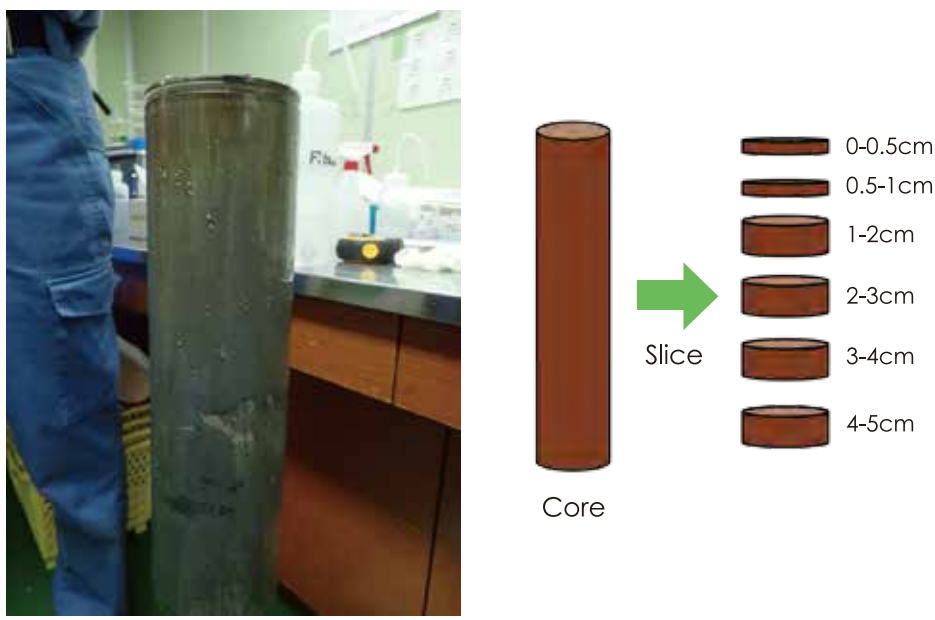


Fig. 2 Example of the core collected and transported inside of a laboratory in a ship and slicing of the core to collect for each layer (right)
Each layer of the core is subsampled according to a layer from a surface layer and its biota is analyzed according to position of the layer

2-2.DNA extraction

Mixed DNAs derived from the organism (microorganism) inhabiting the sample are extracted for amplicon sequence analysis. At this time, it is important to use method that is able to comprehensively recover DNAs derived from all organisms in the sample. Further,

since the samples derived from soils (sediments) often contain impurities such as humic acid, which inhibits PCR, it is desired to use a method which can remove these contaminants from samples.

The DNA extraction process includes cell disruption, removal of cell components such as proteins, and purification of DNAs. Among these steps, the cell disruption step is conducted by physical disruption with beads or chemical disruption with surfactants and cell wall digesting enzymes. Generally, both physical and chemical disruptions are usually performed together. Since cell disruption efficiency will affect the detection rate of organisms at the end of analysis, method selection is important. Further, DNA extraction kits which provide the process described above are currently commercially available from various companies and many reports have been published which compare extraction efficiencies obtained by different DNA extraction kits. During extraction, the amount of sample used, disruption method, and extraction method (kit) should be recorded.

2-3.Construction of amplicons by PCR

PCR is performed using primers which can amplify a target nucleotide sequence from the obtained diverse DNAs. Before performing PCR, the PCR conditions must be determined in order to obtain high quality of amplified products, in addition to selecting primers. The important factors to obtain successful PCR amplicon depend on primes designed, enzyme for PCR, number of PCR cycles, and annealing temperature to be used.

PCR is performed using primers which can amplify a target nucleotide sequence from the obtained diverse DNAs. Before performing PCR, the PCR conditions must be determined in order to obtain high quality of amplified products, in addition to selecting primers. The important factors to obtain successful PCR amplicon depend on primes designed, enzyme for PCR, number of PCR cycles, and annealing temperature to be used.

1 Primer (for amplification of a target gene) selection

The genes which are often used for estimation of biological species are ribosomal RNA (rRNA) genes. Specifically, the nucleotide sequence of small-subunit rRNA is used for analysis of prokaryotes (16S rRNA gene) and eukaryotes (18S rRNA gene) (Fig.3). Since databases are becoming more completed for rRNA gene sequences, it is becoming possible to identify biological species in samples by comparing the obtained sequences with database sequences. Additionally, genes such as cytochrome c oxidase subunit 1 gene (CO1) and cytochrome b gene, which have specific functions, can be used for predicting biological species. Many commonly used genes are registered in databases. Therefore, the number of identifiable biological species will be increased by selecting commonly used genes as analysis objects.

Furthermore, at this moment, there are restrictions on the nucleotide sequence length that can be analyzed by the next generation sequencer. Therefore, it is necessary to use primer pairs that can amplify 400 bases length of sequences, which has sufficient resolution to discriminate each organism, and can comprehensively amplify target genes. The position of the nucleotide sequence recognized by the primers used for the amplification of the 18S rRNA gene for meiofauna analysis is shown in Fig. 4 and the sequences of these primers are shown in Table 1. Approximately 300-500 bp of amplified products are expected to be produced by PCR using primers with the same color.

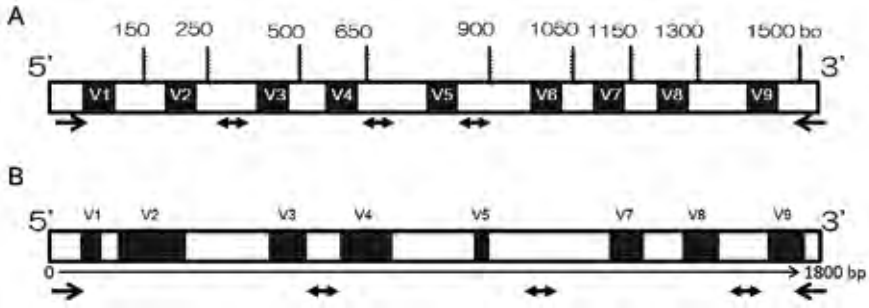


Fig. 3 The schematic diagram of the variable (V1-V9) and conserved region in the small subunit (SSU) rRNA gene sequence. A), Prokaryotes SSU rRNA; B), Eukaryotes SSU rRNA. In Eukaryotic SSU rRNA gene sequence, the V6 region which is corresponding to the prokaryotic SSU rRNAs is more conserved (Neefs et al, 1993). Arrows mean example of universal primer position for PCR amplification.

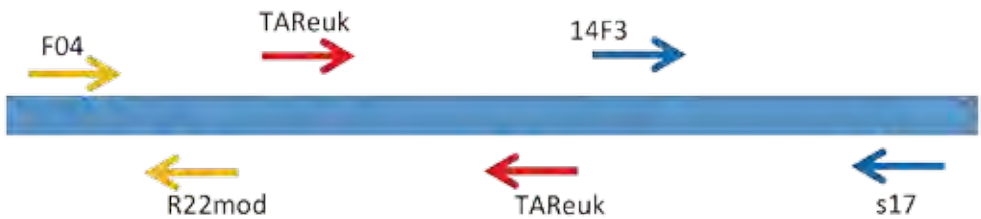


Fig. 4 Location of each primer which targets sequence of 18S rRNA gene Arrows pointing toward 3' from 5' are forward primers while primers pointing opposite direction are reverse primers. In this figure, the range flanked by same color of arrows is amplified.

Table 1. Examples of PCR primer sequences for 18S rRNA gene sequence (First PCR)

Primer	Sequence (5'→3')	Reference
F04-1st	Over hang adaptor sequence-GCTTGTCTCAAAGATTAAGCC	Blaxter,M.L. et al. (1998)
R22mod-1st	Over hang adaptor sequence-CCTGCTGCCTTCCTTRGA	Sinniger F. et al. (2016)
TAReuk454FWD1-1st	Over hang adaptor sequence-CCAGCASCYGCGGTAATTCC	Stoeck, T. et al. (2010)
TAReukREV3-1st	Over hang adaptor sequence-ACTTTCGTTCTTGATYRA	
S14F3-1st	Over hang adaptor sequence-ACGCAMGTGTGAAACTTG	Lejzerowicz F, et al. (2013)
S17-1st	Over hang adaptor sequence-CGGTCACGTTCGTTGC	

* Over hang adaptor sequence is specific sequence for NGS amplicon analysis. The sequence is dependent on used sequencing kit or sequencer.

2 Enzyme for PCR

Various enzymes for PCR are currently commercially available. Examples of enzyme commonly used are, EX Taq, LA Taq (TAKARA), Platinum Taq DNApolymerase (Thermo Fisher Scientific), etc. Requirements of the enzyme include high fidelity, low attenuation (slow) even when the number of cycles increases, and resistance to amplified nucleotide sequences and inhibitors. In order to reduce PCR errors, each sample will be amplified with 3 replicate PCR reactions. The PCR amplifications are generally done in 3 separate PCR tubes, after electrophoresis to check the amplified product, combined triplicate PCR reactions into one PCR tube so that all 3 replicate reactions of the same sample are pooled together. After that the pooled samples will be treated following steps. An example of PCR reaction mixture is shown in Table 2. Please record type of enzyme used.

Table 2 An Example of PCR reaction mixture

	Amount/ tube (μL)
10×Buffer	2
dNTP	1.6
Forward primer (10 μM)	0.4
Reverse primer (10 μM)	0.4
Ex Taq (Takara)	0.1
Template DNA(0.5–1 ng/uL)	1
DW	14.5
Total volume	20

3 PCR — cycle condition Confirmation of PCR amplified products

In PCR reaction, amplified products are constructed by repeating a cycle (thermal cycle), which consists of 3 steps, a denaturation step (denaturing a double stranded DNA into a single stranded DNA by heat), an annealing step (annealing of a primer with a single stranded complementary DNA), and an elongation step (elongation of a single stranded DNA from the site where a primer annealed) for dozens of times, (Fig. 5). A PCR machine which operates the series of reaction has been commercially available (Fig. 6). An example of PCR cycle condition is shown in Table 3. As the number of the cycle increases, the amount of amplified products as well as mistakes increase; therefore, it is desired that the number of cycles should be minimized in PCR reaction. Thus, it is necessary to determine the condition which produces adequate amplified products in amplicon sequence analysis. Furthermore, annealing temperature of the primers with nucleotide sequence is an important factor for PCR where high annealing temperature causes less misannealing of primers although it causes failure in amplification while lower annealing temperature causes nonspecific amplification[7]; therefore, it is necessary to determine annealing temperature for each condition. Instrument type, cycle number, temperature, time, etc. should be recorded for PCR.

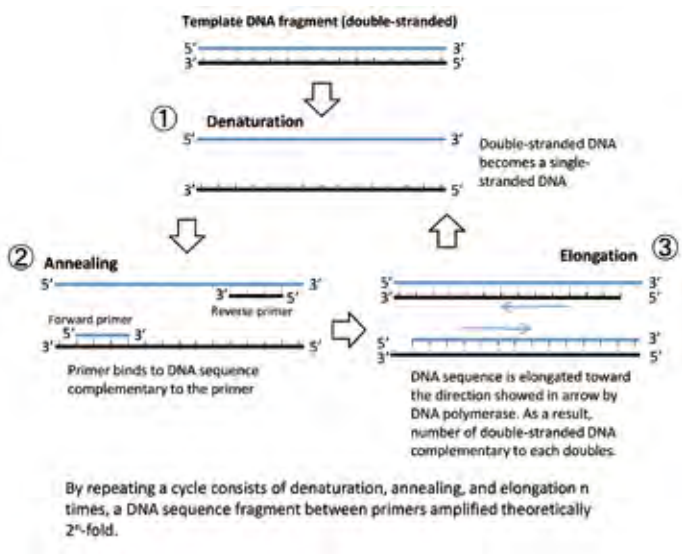


Fig. 5 Outline of PCR (Nishijima, 2013)

Fig. 6 Example of PCR instrument

Table 3 Example of PCR Reaction Conditions

Temperature (°C)	Time	Number of Cycles
96	1 min	1
96	25 sec	20-30*
55	45 sec	
72	1 min	
72	7 min	1
4	∞	

* The number of cycles can be changed to the number of times that an appropriate amplification product can be obtained.

4 Confirmation of PCR amplified products

PCR amplified products are confirmed by agarose gel electrophoresis. Agarose gel electrophoresis can be used to examine whether or not a target fragment is amplified, to estimate approximate amount of a target fragment produced by intensity of the fragment band without quantitation, or to examine whether or not non-specific amplified products are present (Fig. 7). Commonly agarose concentration of 1~2% dissolved in TAE or TBE buffer is used. High purity agarose for DNA analysis should be used for electrophoresis. Moreover, a DNA marker with an appropriate range should be electrophoresed simultaneously in order to confirm the length of amplified fragment (Fig. 7, M).

An example of electrophoresis machines includes Mupid. Several dyes are available on market which enable to visualize PCR amplified fragments. Ethidium bromide, which is carcinogenic, had been previously used; however, it is replaced with dyes which has higher safety, can be handle easily, and has higher sensitivity. These dyes include RedSafe which can be added to agarose gel beforehand at the time of agarose gel preparation, EzVision which is mixed to samples at the time of electrophoresis, and SYBR Green and SYBR Gold which are used as staining solution of a gel after electrophoresis is completed. Gel Red and GelGreen which can be used either by adding to a gel prior to electrophoresis or by using as staining solution of a gel after electrophoresis are also available on market. Sensitivity of detection of fragments depends on dyes; therefore, it is recommended to use dyes with higher sensitivity in order to identify presence of nonspecific PCR products as much as possible. Agarose concentration, electrophoresis condition such as buffer, voltage, time (min), and applied amount should be recorded.

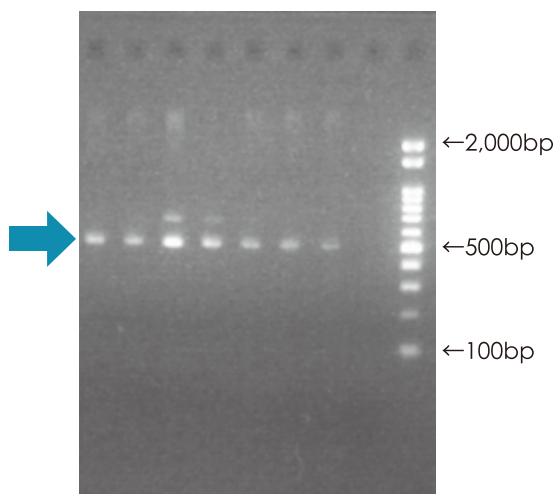


Fig. 7 Example of agarose gel electrophoresis

Expected amplified fragment (app. 500bp) is detected in lane 1~7 (blue arrow). An extra band above a 500bp band is observed in samples in lane 3 and 4. M indicates a size marker.

5 Purification of PCR amplified products

After confirming the presence of an appropriate size of a PCR amplified product by agarose gel electrophoresis, unnecessary primers and reagents are removed for the use of the PCR amplified product for the next step.

Although various kits are provided by manufactures, a kit such as ExoSAP-IT® (AFFYmetrix) is convenient. Methods used should be recorded.

6 Index PCR

Performance of next generation sequencing depends on the type of a sequencer; however, it is known to read 10 million sequences (called a read) by one analysis. Thus, it is possible to analyze multiple samples simultaneously and it is often used to sequence all of mixed multiple PCR amplified products produced from an environmental sample. However, the number of a PCR amplified products in each sample decreases in case where excess number of samples are analyzed at one time; therefore, it is desired to use sample number of less than 100 at each analysis . To differentiate samples in a mixture, marker nucleotides are attached to each sample by PCR. This is called Index PCR. In Index PCR, primers with unique nucleotide sequences (ID tags) for each sample are used for a second PCR which uses PCR amplified products obtained in the first PCR as templates to attach ID tags to the PCR amplified products (Fig.8). It becomes possible to differentiate sequences by sorting them according to their identification tags. ID tags can be added using a commercially available kit (e.g.: Nextera XT Index Kit, Illumina). Following PCR, agarose electrophoresis is performed according to procedures as described in (4) and amplified fragments with predicted length are confirmed.

Sequences of tags used for Index and samples as well as the PCR conditions (enzyme, cycle, etc.) should be recorded. Furthermore, it is also desired to use a minimum number of cycles in Index PCR.

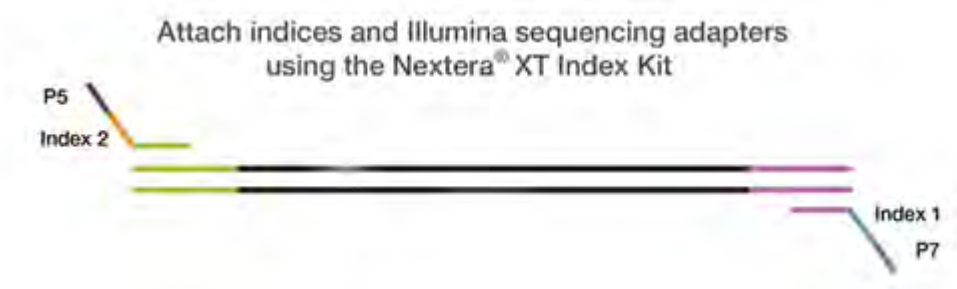


Fig. 8 Location of index tag sequences (Index 1 and 2) attached by Index PCR Example of Miseq, Illumina (16s-metagenomic-library-prep-guide-15044223-b).
The sequences of index 1 and 2 are appended after each direction of the overhang adapter sequence. Regions in black indicate nucleotide sequence to be analyzed while regions in purple and yellow green indicate primers of each gene which contain sequences of overhang adaptors necessary for analysis by next generation sequencer.

7 Purification of Index PCR amplified products using magnetic beads

Amplified PCR products which showed expected fragment length by electrophoresis following Index PCR are purified using magnetic beads (AMPure XP: Beckman Coulter), etc. to remove residual PCR primers and PCR reagents. During purification, the procedure should be conducted carefully so that beads are not excessively dry and beads are not suctioned and discarded.

8 Concentration measurement by quantitative PCR

The concentration of each amplified product following Index PCR is different; therefore, nucleotide sequences of amplified products with higher concentration are sequenced more often than those with lower concentration when samples are mixed for analysis by next generation sequencer. To avoid this, quantitative PCR is performed to measure concentration of each amplified product. Concentrations of each amplified products are adjusted so that all products are mixed at equal concentrations. For quantitative PCR, when Miseq from Illumina is used as a next generation sequencer, for example, a specialized kit is commercially available (e. g.; KAPA Library Quantification Kits for Next-Generation Sequencing, KAPA Biosystems). The instrument used for quantitative PCR and methods, etc. should be recorded.

9 Quality check of amplified products by a bioanalyzer

Fragment lengths of samples purified by magnetic beads and impurities contained in amplified products (including products with unpredicted fragment lengths) are examined by Bioanalyzer, Agilent 2100. Quality of samples which are to be analyzed by next generation sequencer is checked by this process.

2-4.Sequence

Sequencing of samples where each sample is adjusted to have the same concentration are performed by next generation sequencer. Sequencing methods are different depending on the manufactureres and type of instruments. The most commonly used sequencer is Miseq (Fig. 9) or Hiseq which is a high-end instrument developed by Illumina. The instrument used and number of samples analyzed as well as the concentration of each internal standard and analysis subject should be recorded.

2-5.Data analysis

Sequences generated by a sequencer are sorted according to an identification tag attached using Index PCR and nonspecific DNA fragments determined by length of nucleotides are removed (removal of low quality reads). Generated sequences are sorted into a classification unit called Operational Taxonomic Unit (OTU) according to the similarity of the sequences, and the classification group and composition ratio are estimated in the sample using OTU as a unit. Softwares and database used for data analysis should be recorded.

Fig.9 Next generation sequencer, Illumina Miseq
(http://jp.illumina.com/systems/miseq.html)

Notes

*1. Microbes in this report include not only bacteria, fungi and yeasts but also protist, minute animal and plant planktons, and nematodes.

*2. DNA extraction

During DNA extraction, intracellular DNA is eluted by physical or chemical disruption of cells, purified, and recovered. During the extraction, if there is a difference in disruption or dissolution of cells, it means that DNAs derived from organisms that have difficult-to-break cells cannot be fully recovered. Therefore, population of organisms with such cells can be estimated lower than actual.

*3. PCR amplification

A copy number of a target gene differs depending on biological species and amplification degree depending on conditions used in the PCR cycle such as primers, enzymes, annealing temperature, and a number of a cycle. PCR amplification efficiency indicates these differences.

*4. PCR

PCR is an abbreviation for Polymerase Chain Reaction, DNA sequence synthesis reaction is repeated to amplify DNA fragments of target regions of hundreds of thousands of times of genes using template DNA, and two types of short DNA fragments, and DNA polymerases.

*5. Cloning

Conventional sequence analysis called the Sanger method cannot be used for analysis of PCR amplified products generated from diverse DNAs. Therefore, these amplified DNAs are first cloned into a plasmid derived from E.coli for purification process. By culturing E,coli containing these plasmids, number of plasmids increases resulted in increased amount of PCR amplified DNA sequences. These plasmids are then analyzed by the Sanger method after extraction and purification of plasmids. The process is complex as well as costly and time consuming due to the requirement of cloning and culturing during the procedure; therefore, the Sanger method is considered to be unsuitable to comprehensively analyze biological species in environment.

*6. DGGE

DGGE is an abbreviation for Denaturing Gradient Gel Electrophoresis and is a method to separate PCR amplified nucleotides of 16S rRNA genes extracted from diverse DNAs in an environmental sample and analyze the separated nucleotide sequences. The method has been introduced by Muzer et al. using bacterial community as a target for analysis (Muyzer et al., 1993). The method can be applied to organisms such as fungi and nematodes by changing target genes (Möhlenhoff et al., 2001; Waite et al., 2003).

*7. Nonspecific amplification

Amplified genes and regions are determined by the primer pairs in PCR. However, PCR amplification may occur despite the difference in the primer sequences. For example, when there are many complementary portions in the sequences of the primers, the sequence in which the primers are bound to each other is amplified. These unintentionally produced amplified products are called nonspecific amplified products.

*8. It is possible to predict number of entire species living in a specific environment by constructing rarefaction curves based on the read number and the OTU number obtained by data analysis. It is thought to be desirable that the number of reads which can be analyzed should be kept large; therefore, it is recommended that the number of samples used for each analysis should be below 100.

References

Blaxter,M.L. et al. (1998) A molecular evolutionary framework for the phylum Nematoda. Nature, 392, 71–75.

Illumina 16s-metagenomic-library-prep-guide-15044223-b: https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf Accessed 24 Feb. 2017.

International Seabed Authority. (2013). Recommendations for the guidance of contractors for the assessment of the possible environmental impacts arising from exploration for marine minerals in the area. https://www.isa.org.jm/sites/default/files/files/documents/isba-19ltc-8_0.pdf. Accessed 24 Feb. 2017

Dahlgren T (2014). Advances in molecular methodologies and the application and Interpretation of molecular methodologies for macrofauna classification – Their relevance to environment assessment and monitoring. In: International Seabed Authority. https://www.isa.org.jm/files/documents/EN/Workshops/2014b/Abs/Dahlgren.pdf Accessed 24 Feb. 2017

Sinniger F et al. (2016). World wide analysis of sedimentary DNA reveals major gaps in taxonomic knowledge of deep-sea benthos. Front. Mar. Sci. 3:92. doi:10.3389/fmars.2016.00092

Lejzerowicz F, et al. (2013). Ancient DNA complements microfossil record in deep-sea subsurface sediments. Biol Lett 9: 20130283.

Muyzer G.et al. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59, pp.695–700

Möhlenhoff P. et al (2001). Molecular approach to the characterisation of fungal communities: methods for DNA extraction, PCR ampli ¢ cation and DGGE analysis of painted art objects. FEMS Microbiol Lett 195, pp.169–173.

Nishijima M. (2013). Idenshi kaiseki-hō no genre. In Guidebook of Easy and Rapid Microbial Test Methods. Ed. Igimi, S., Ezaki, T., Takatori, K. and Tsuchido, T. pp. 43-47. Technosystem, Tokyo. [in Japanese]

Neefs J-M. et al (1993) Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res 21:3025–3049.

Stoeck, T. et al. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol. Ecol. 19 suppl1, 21–31.

Waite IS et al. (2003). Design and evaluation of nematode 18S rDNA primers for PCR and denaturing gradient gel electrophoresis (DGGE) of soil community DNA. Soil Biol Biochem 35, pp.1165–1173.

Sample

Sheet No. _____

Metagenome analysis	Sample
Sample	Original sample No.
Sample ID No.	No.
Sampling date (dd/mm/yyyy)	

Sample	Sediment	Water	Algae	Animal	Others

Location (Point; Latitude, Longitude, Depth)	Latitude	Longitude	Depth (m)

Sampling method	Core sampler	Niskin

Sub Sampling	Core	Filtration/Sieve	Layer (cm)

Biological sample	Number of individuals	region	Weight	Size

Preservation					
Stock solution	LifeGuard	RNAlater	Guanidine	TE	Folmaldehyde
Temperature	-80℃	-30℃	-20℃	4℃	RT

Note						
	TOC	Pore water	PH	Temperature	Morphological observation (Microscope; FlowCam)	Bacterial cell count
Other analyses:						
Temperature						

Nucleic Acid Extraction

Sheet No. _____

Metagenome analysis	Extraction for nucleic acid	
Sample		
Sample ID No.	No. _____	
Sampling date (dd/mm/yyyy)		
Sample usage	Weight	Volume

Nucleic acid extraction						
Target	DNA		RNA		Co-extraction	
Kit	MO BIO	Isoil	MO BIO	Isoil	MO BIO	Isoil
Others / Note						

Purification for extracted nucleic acid	DNA	RNA
method	Rnase	Dnase

Extracted nucleic acid				
measurement for concentration and purity	Spectrophotometer		Instrument :	
	230	260	280	nm
Concentration	Qubit		NanoDrop	
Fragmentation check	Agarose gel electrophoresis :			

Amplicon construction

Sheet No. _____

Metagenome analysis	Amplicon sequence analysis
Sample	Original sample No.
Sample ID No.	No.
Date (dd/mm/yyyy)	

Amplicon sequence analysis

Target gene (region)	16S	18S	CO1	cyt	Other
Region	**V3-V4**	**V1-2**	**V9**	**f37**	**Other**

Primer	Forward		Reverse		
Prokaryotes	341F	530f		534R	907r
Eukaryotes	F04	TAReuk454 FWD1	S14F3	R22mod:	TAReukREV3	S17
Other						

cDNA library (reverse trascription) (RNA→cDNA)

Primer	Target gene (region)	Reverse

PCR condition	Amplicon check
Sterile distilled water	**Without template RNA**

PCR

Enzyme (Taq)	LA Taq	ExTaq	MightyAmp	Other

Cycle

Instrument	ABI	Takara	Other

Electrophoresis	Conditions：Agarose gel %, Buffer： , time (min)

Index PCR	Tag

Purification	Purification Kit	Magnet beads
Measurement for concentration	**Quantitative PCR**	**Instrument: ; Kit**

Fragmentation check	BioAnalyzer

Sheet No. _____

Metagenome analysis	Next generation sequence
Sample	
Sample ID No.	No.
Date (dd/mm/yyyy)	

Next generation sequencer				
Instrument	Illumina Miseq	ABI IonPGM		
Kit	MiSeq Reagent Kit	ABI		
Concentration adujstment of amplicon				
nm/ sample				
"Internal standard mixing ratio (Illumina Miseq) "	1 : 1			

Memo

Data Analysis

Sheet No. _____

Metagenome analysis	Data analysis
Sample	Original sample No.
Sample ID No.	No.
Date (dd/mm/yyyy)	

Sample	number of obraitned reads		
Removale reads	**Number of low quality reads**	**Chimeric sequence**	**Primer sequence**

Assembled of paired-end sequence	fastq-join		

Software for analysis	Qiime	Megan	mothur
Other software			

OTU	Number of OTU	BLAST
Similarity (%)		
Other	**Similarity indices**	

Memo

Shotgun Sequence Analysis

Sheet No. _____

Metagenome analysis	Next generation sequence analysis (Shot gun sequence analysis)
Sample	
Sample ID No.	No.
Date (dd/mm/yyyy)	

Next generation sequencer		
Instrument	Illumina Miseq	ABI IonPGM
Reagent	MiSeq Reagent Kit	ABI
Fragmentation	Enzyme	Physical Fragmentation
	Kit:	Instrument:
Attachement of tag sequence for differentiation of a sample		Enzyme
		Kit:
DNA concentration adjustment		
"Internal standard mixing ratio (Illumina Miseq) "		1：1
p		
Purification		
Magnetic beads		

Memo

Memo

Memo

SIP Protocol Series

SIP Protocol No.1

Application of environmental metagenomic analyses for environmental impact assessments

SIP Protocol No.2

Genetic Connectivity Survey Manuals

SIP Protocol No.3

A rapid method to analyze meiofaunal assemblages using an Imaging Flow Cytometer

SIP Protocol No.4

Acquisition of Long-Term Monitoring Images Near the Deep Seafloor by Edokko Mark I

SIP Protocol No.5

Microstructure Measurements aroud Deep Sea floor

-Direct Measurements of the Deep Sea Turbulence flow-

Contact

JAMSTEC (Japan Agency for Marine-Earth Science and Technology)
2-15, Natsushima-cho, Yokosuka-city, Kanagawa, 237-0061, Japan

TEL +81-46-866-3811 FAX +81-46-867-9755

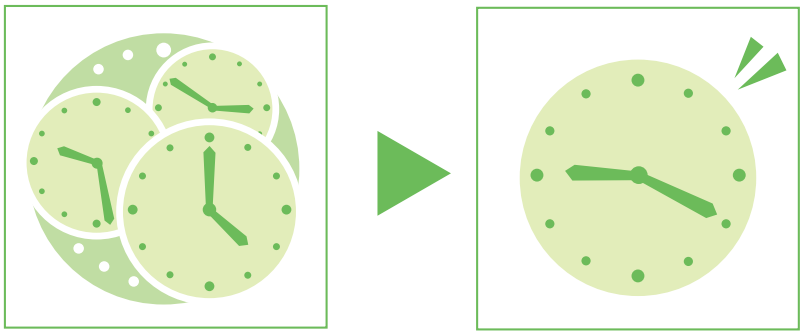
E-mail sip-pc@jamstec.go.jp

ISBN 978-4-901833-27-1

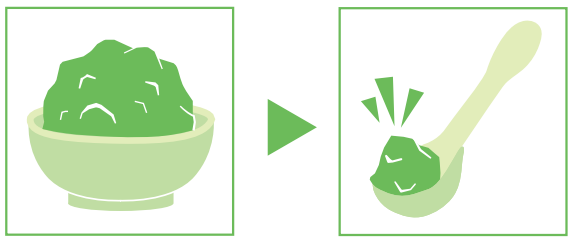
Version.1
May.1.2017

before

after

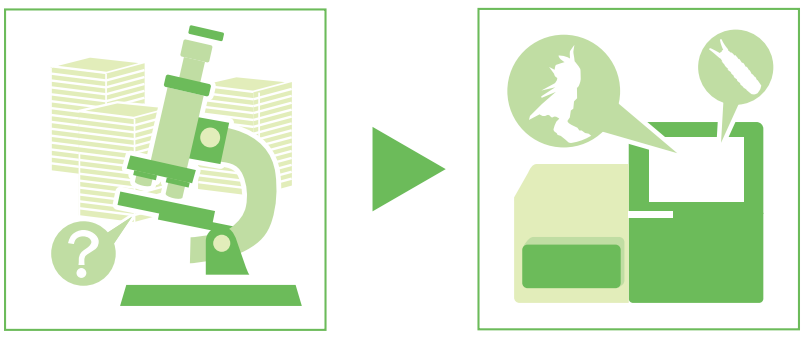


Speedy analysis possible

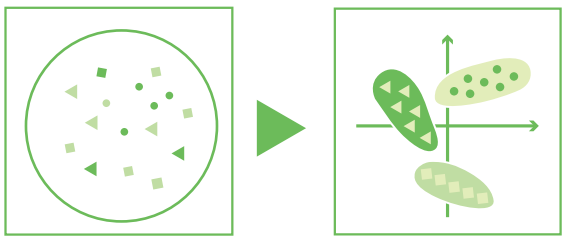


Only small amounts of sample necessary

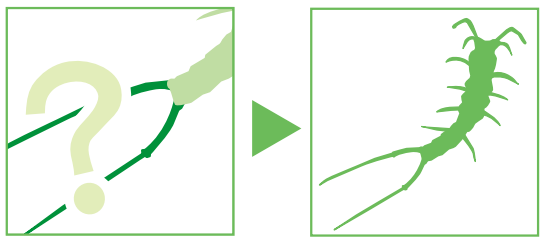
Even if the species' names are not resolved, the nucleotide sequences can be compared to existing databases for identification.



Objective and comprehensive data acquisition possible



From the acquired genetic information, inter-regional comparisons of biological communities become possible.



DNA information enables identification of samples in early developmental stages (egg, larva and immature specimen) and for damaged samples.

JAMSTEC 国立研究開発法人 海洋研究開発機構
JAPAN AGENCY FOR MARINE-EARTH SCIENCE AND TECHNOLOGY